



ELSEVIER

Contents lists available at ScienceDirect

# Infant Behavior and Development

journal homepage: [www.elsevier.com](http://www.elsevier.com)

Full length article

## Vocal responsiveness of preterm infants to maternal infant-directed speaking and singing during skin-to-skin contact (Kangaroo Care) in the NICU

Maria Eduarda S. Carvalho <sup>a,\*</sup>, João M.R.M. Justo <sup>b</sup>, Maya Gratier <sup>c</sup>, Teresa Tomé <sup>d</sup>, Esmeralda Pereira <sup>d</sup>, Helena Rodrigues <sup>a</sup>

<sup>a</sup> CESEM-NOVA-FCSH, Portugal

<sup>b</sup> Faculdade de Psicologia, Universidade de Lisboa, Portugal

<sup>c</sup> Laboratoire Ethologie, Cognition, Développement, Université Paris Nanterre, France

<sup>d</sup> Maternidade Dr. Alfredo da Costa – Centro Hospitalar Universitário de Lisboa Central, Portugal

### ARTICLE INFO

#### Keywords:

Preterm dyads  
Maternal speaking  
Maternal singing  
Preterm infants' vocalizations  
Expressive timing  
Kangaroo care

### ABSTRACT

Vocalizations of full-term newborns occur in a short latency time during the neonatal period. Contingent response time of preterm babies is still unknown. An increase of preterm babies' vocalizations following exposure to parental speech was also observed. Mothers and babies co-modulate their vocalizations in preterm dyads. Purpose: To observe temporal features of maternal and infants' vocalizations in speaking and singing conditions in preterm dyads. Methods: In a NICU mothers (N = 36) were invited to speak and to sing to their preterm infants during Kangaroo Care. Microanalysis of temporal units were performed with ELAN Software. Results and conclusions: Preterm infants vocalize less often while their mothers speak and sing than during baseline and their vocalizations tend to be more alternating in the speaking condition and more overlapping in the singing condition. It is also concluded that preterm infants take more time to respond to maternal speaking than to maternal singing.

### 1. Introduction

The ontogeny of human communication in infancy is the main focus of this paper (Bateson, 1979; Bruner, 1983; Maturana, 1978). The temporality of the human communication in infant-directed speech is one of the aspects underlined in this field (Gratier & Trevarthen, 2008; Malloch & Trevarthen, 2009).

Term newborns' preference for the mother's voice compared to an unfamiliar female voice (DeCasper & Fifer, 1980) or to the father's voice (Lee & Kisilevsky, 2014), provides evidence of an experience-based prenatal auditory memory (DeCasper & Spence, 1986; DeCasper, Lecanuet, Busnel, Granier-Deferre, & Maugeais, 1994). Fetal cardiac response to the mother's voice has been observed at 32–34 weeks of gestation (Kisilevsky & Hains, 2011) and activation of the lower bank of the left temporal lobe has been observed through fMRI measures at 34 weeks of gestation (Jardri et al., 2012). Marx and Nagy (2015) found that, as in newborns, maternal voice elicits behavioral attentional-orientation responsiveness in the fetus. In their study, decrease of fetal yawning was considered an indicator of fetal attentional response to maternal voice when compared with maternal touch. Fetal mouth opening was also found to index discrimination between maternal singing and non-communicative maternal stimuli (Ferrari et al., 2016).

\* Corresponding author.

Email address: [educarte@sapo.pt](mailto:educarte@sapo.pt) (M.E.S. Carvalho)

Furthermore, auditory discrimination of temporal and spectral features between linguistic stimuli and musical melody has been found in fetuses (Granier-Deferre, Ribeiro, Jacquet, & Bassereau, 2011; Jardri et al., 2012). Also, in one-day-old newborns, discrimination between different linguistic stimuli (native vs. non-native language stories) as well as discrimination between different non-linguistic stimuli (native vs. non-native stimuli played in reverse) were observed (Vannasing et al., 2016).

A few weeks after birth, infants' vocalizations are described as coos and murmurs (Oller, 2000). These vocalizations are usually associated with an intentional knit-brow gaze directed to the social partner together with mouth movements similar to speech movements that Trevarthen (1993) designated as pre-speech movements. Mothers respond more frequently to babies' vocalizations than to change in their gaze orientation or to smiles (Van Egeren, Barratt, & Roach, 2001).

Maternal responsiveness prompts the baby to respond reciprocally (Papousek, Papousek, & Bornstein, 1985; Stern, 1985; Stern, 1999). Mothers have been found to contingently respond to newborn vocalizations within a 1000 ms temporal window and newborns in turn respond to mothers within a latency time of 1000 ms giving rise to turn-taking sequences (Dominguez, Devouche, Apter, & Gratier, 2016). Contingent maternal behavior can increase the frequency of infants' vocalizations (Goldstein & Schwade, 2008; Pelaez, Virués-Ortega, & Gewirtz, 2011). It has been shown that preterm infants in the NICU also vocalize more often when adults', including their parents, speak in their vicinity (Caskey, Stephens, Tucker, & Vohr, 2011). This finding supports the hypothesis that turn-taking is a precocious human ability (Levinson, 2016). Most research to date has shown that infants' sensitivity to social contingency develops at around two months of age (Striano, Henning, & Stahl, 2005) when maternal and infant vocalizations alternate with pauses varying from 500 ms to 1000 ms (Jaffe, Beebe, Feldstein, Crown, & Jasnow, 2001). The response latency time of maternal utterances to babies of 2–4 months of age is less than 1000 ms (Keller, Lohaus, Völker, Cappenberg, & Chasiotis, 1999). When babies are 4 months old, a period of 3000 ms has been defined as the most frequent contingency window for social stimuli (Van Egeren et al., 2001).

In the neonatal period, newborns' vocalizations are modulated by maternal vocalizations (Rosenthal, 1982). The study with newborns by Dominguez et al. (2016) found that overlapping vocalizations totaled one third of all babies' vocalizations. According to Lavelli and Fogel (2013) during the earliest months of human life vocal interaction is bidirectional and regulated by both the mother and the infant. Ginsburg and Kilbourne (1988) found more overlapping vocalizations in interactions between mothers and 7–13 week-olds compared to those between mothers and 12–18 week-old infants. Another longitudinal study on vocal turn-taking by Hilbrink, Gattis, and Levinson (2015) showed a decrease in overlapping vocalizations after five months of age and found that pauses between vocalizations were longer around the age of nine months. In yet another longitudinal study, Gratier et al. (2015) found that pauses between mothers' and infants' vocalizations were longer on average for infants aged between 17 and 21 weeks than for infants aged between 8 and 13 weeks, while overlapping vocalizations of younger infants were longer than those produced by older infants. Also the proportion of maternal overlapping vocalizations was much higher in older dyads than in younger dyads. These studies suggest that infants acquire turn-taking ability progressively.

Comparing preterm dyads two months after discharge (37–40 weeks GA) with dyads with 2-month-old term infants, it was observed that mothers of preterm infants vocalized contingently more frequently than mothers of term infants, while term infants vocalized contingently more frequently in relation to their mothers' vocalizations than preterm infants (Reissland & Stephenson, 1999). Salerni, Suttora, and D'Odorico (2007) also observed that preterm mothers are more responsive than mothers of term infants.

Furthermore, preterm neonates in the NICU, when in a stable medical condition, respond to both maternal infant-directed speech and singing. Filippa, Devouche, Arioni, Imberty, and Gratier (2013) found changes in heart-rate and oxygen saturation during maternal vocal contact with preterm infants in their incubators and found that infant-directed speech was associated with more awake behavioral states whereas infant-directed singing was associated with more drowsy or active sleep states. In preterm dyads, an increase in maternal pitch during both the speaking and singing conditions was found following positive responses to maternal vocalizations (Filippa et al., 2013).

Discrimination between maternal speech and maternal singing has been studied in babies older than six months showing more visual attention and less motor activity in relation to infant-directed singing (Nakata & Trehub, 2004). The ability to discriminate between speech and singing was observed using different vocal responses in babies aged between 9 to 11 months (Reigado, Rocha, & Rodrigues, 2011) as well as in babies between 12 and 24 months (Reigado & Rodrigues, 2017). During the second year of life, babies' vocal responses seem to be more frequent in a speaking condition than in a singing condition (Reigado & Rodrigues, 2017). Interestingly, female babies were found to produce more vocalizations than male babies in the speech condition while in the singing condition no differences were found.

Newborns are able to identify temporal units in language based in rhythmic and intonational cues (Byers-Heinlein, Burns, & Werker, 2010; May & Byers-Heinlein, 2011; Moon, Lagercrantz, & Kuhl, 2013; Ramus, Nespor, & Mehler, 1999) and young infants are sensitive to the phrase-final lengthening in infant-directed speech (Trainor & Adams, 2000). Both infant-direct speech and singing are characterized by temporal units that can be hierarchically expressed at multiple time scales from hundreds of milliseconds to tens of seconds (Falk & Kello, 2017). When compared with adult-directed speech, more temporal clustering is observed in infant-direct-speech. This is especially remarkable in the variability of durations of phrase-final lengthening and also in variability of duration at the phrasal constituent level (Falk & Kello, 2017). These observations suggest that utterance-final variations play a fundamental function for babies' perception of auditory messages. As Delavenne, Gratier, and Devouche (2013) found, infant-directed singing is probably characterized by a longer final utterance, when compared with the length of sentences' final utterances during infant-directed speech; temporal structure in infant-direct-singing is also indexed by phrase-final lengthening and by greater variability at the pulse level. Interestingly, maternal singing directed to female infants presents a higher phrase-final lengthening than singing directed to male infants (Delavenne et al., 2013).

Maternal lullabies have been found to improve maternal-infant bonding as well as to benefit neonatal behavior and maternal stress (Persico et al., 2017). A recent study shows that maternal singing is connected with maternal perception of greater emotional closeness to the baby as well as with an increase of positive affect, a decrease of negative affect and a decrease of maternal anxiety (Fancourt & Perkins, 2018).

Maternal singing during skin-to-skin contact has beneficial effects on autonomic stability in preterm babies (32–36 gestational weeks) as also plays a part in the reduction of maternal anxiety levels (Arnon et al., 2014). A recent study on the use of skin-to-skin contact between mothers and preterm infants aged between 25 and 32 gestational weeks placing the baby in a diagonal position — known as Flexion Diagonal Soutenué (FDS) or kangaroo supported diagonal flexion positioning — favoring visual contact has shown that mothers display more behaviors of attention directed to signals from the babies, as well as more smiles and more contingent vocalizations (Buil et al., 2016). Thus kangaroo supported diagonal positioning is thought to be more favorable and natural for the mother to speak and to sing for her preterm baby in a contingent way. A recent systematic review of maternal voice interventions in preterm infants (Filippa et al., 2017) concluded that more studies using contingent maternal speech and singing are needed in neonatal care units. Few studies have dealt with the timing structure of maternal infant-directed speech to preterm infants in the neonatal period. And no study to date has investigated the timing structure of maternal infant-directed singing to preterm infants in the neonatal period. In spite of the existence of observations of preterm babies' vocalizations in the presence of parents' voices (Caskey et al., 2011) it is important to study the expressive timing of preterm dyads' vocalizations in both the speaking and singing conditions. Given what we have learned about the preterm infants' vocal responsiveness to speech in the NICU and its connection with language ability later on (Best, Bogossian, & New, 2018; Caskey et al., 2011; Caskey, Stephens, Tucker, & Vohr, 2014), it seems worthwhile to look more closely at vocal interaction between mothers and infants in the NICU. Specifically, the timing patterns of infant-directed speech and singing to preterm infants merit attention as well as the timing of infants' vocal responsiveness to speaking and singing.

Thus, one of the main goals of this study is to enhance our knowledge of the specificity of the temporal structure of maternal speech and singing directed to the preterm baby in skin-to-skin and face-to-face situation (FDS condition) in the neonatal care unit. Another goal is to identify the timing and durational features of preterm babies' vocalizations in these two conditions. This study aims to clarify aspects related to the role of maternal speech and maternal singing in the development of preterm infants in the context of neonatal care units. Finally we aim to understand the differentiation of the expressive timing of maternal speech and singing and the implications of this in the development of preterm babies' behavior in the context of NICUs.

## 2. Methods and materials

### 2.1. Inclusion and exclusion criteria

Fifty mothers of preterm infants accepted to participate in this study and signed an informed consent form after being advised about the goals and procedures of the study. For clinical and circumstantial reasons 10 dyads were lost. Another 4 dyads were excluded from the study due to perturbations during the recording procedure. Among the observed dyads ( $n = 40$ ), 36 dyads were selected for the study. Recruitment was performed between June 2016 and April 2017 at the Intermediate Neonatal Care Unit of the hospital Maternidade Dr. Alfredo da Costa at Lisbon, after approval by the Ethical Committee of the Central Lisbon Hospital Center (267/2015). Participants were excluded from the study based on the following criteria: a) mother being younger than 19 years old; b) difficulties understanding and speaking the Portuguese language, c) auditory deficit in infant or mother, d) gestation without medical supervision, e) previous psychiatric pathology or serious negative emotional states and f) addictive behaviors. Dyads were also excluded if, at the identification moment, babies had: a) post-menstrual age lesser than 32 weeks or higher than 37 weeks, b) instability of the vital parameters, c) support of CIPAP (Continuous Positive Airway Pressure), d) intraventricular hemorrhages and e) congenital or neurological anomalies of auditory cortex, e) nasogastric tube and f) supporting breathing. Dyads were also excluded if skin-to-skin kangaroo care had not been practiced at least once.

### 2.2. Sample characteristics

Mothers' mean age was 34 ( $SD = 5.63$ ,  $min. = 21$ ,  $max. = 48$ ), most of them were Portuguese nationals ( $n = 26$ , 72.2%), with a high number of successful years at school ( $M = 15.33$ ,  $SD = 3.69$ ,  $min. = 6$ ,  $max. = 24$ ), married ( $n = 23$ , 63.9%; single = 19.4%; out of wed lock = 16.7%) and primiparous ( $n = 22$ , 61.1%). Among the infants, 20 (55.56%) were boys and 16 (44.44%) were girls, their mean gestational age at birth was 30 weeks and 4 days ( $M = 212.78$ ,  $SD = 17.11$ ,  $min. = 178$ ,  $max. = 241$ ) and 34 weeks and 1 day of post-menstrual age at observation ( $M = 239.27$ ,  $SD = 9.28$ ,  $min. = 224$ ,  $max. = 259$ ). Their mean chronological age was 26.5 days ( $M = 26.5\%$ ,  $SD = 19.99\%$ ,  $min. = 4$ ,  $max. = 81$ ), mean weight at birth was 1265.47 g ( $SD = 308.20$ ,  $min. = 590$ ,  $max. = 2017$ ) and mean weight at observation was 1538.05 g ( $SD = 237.72$ ,  $min. = 1060$ ,  $max. = 2185$ ).

About clinical variables, the vast majority of mothers in our sample had no spontaneous abortions ( $n = 23$ , 63.88%), no voluntary abortions ( $n = 30$ , 83.33%) and no abortions for medical reasons ( $n = 23$ , 63.88%). In most cases, their pregnancy was a singleton one ( $n = 26$ , 72.2%) and in some cases two fetuses ( $n = 7$ , 19.4%) and three fetuses ( $n = 3$ , 8.3%) were observed. Most part of mothers said their pregnancies had been planned ( $n = 31$ , 86.1%), they always desired and the majority of deliveries were caesarean ( $n = 24$ , 66.7%). Labor was spontaneously triggered in 19 cases (55.9%) and induced in 15 cases (44.1%).

### 2.3. Instruments

A Sociodemographic and Clinical Questionnaire was designed for this investigation with items about age, education, citizenship, marital status, occupational status, socio-economic status, number of previous children, obstetrical data (spontaneous abortions, voluntary abortions, interruptions by medical advice, number of fetuses at the last pregnancy) and pediatric data (gestational age at birth, birth weight, delivery type, Apgar, age at observation, weight at the observation).

### 2.4. Equipment and procedure

One camera (Panasonic 4K HC-VX870) was oriented toward the mother and the baby. The camera was connected to an external microphone placed near the mother and baby. The mother was provided with a large scarf for holding the baby in skin-to-skin contact during kangaroo care. According to FDS kangaroo care procedure an additional strap was provided to support the baby's neck in such a way as to enable the mother to see the baby's face.

### 2.5. Design

The research protocol lasted for 15 min with the baby and the mother in a skin-to-skin condition (FDS). To begin with, it was required that the baby should reach a state of quiet alertness or drowsiness (Brazelton & Nugent, 2011, pp. 49–51). After that, during an initial baseline period lasting 3 min mothers were asked to remain silent. Each mother was then invited to sing or speak to the baby also for a period of 3 min, at the end of which an interruption was requested, also lasting 3 min (offset period). After the 3 min offset, participants were asked to either speak or to sing to their babies once again, alternating with the previous speaking or singing condition. At the end, a final 3-min-period of silence was requested (offset period). For the speaking condition, mothers were asked to speak to their babies as they usually did, with no pre-established script. For the singing condition, mothers were asked to hum for their babies improvising an original melody without words. In group 1, mothers began with the singing condition and in group 2 mothers began with the speaking condition.

### 2.6. Coding of vocal behavior

For each dyad, one video recording (MP4) and its corresponding audio record (WAV) were coded. ELAN software (EUDICO Linguistic Annotator) version 4.9.4 was used for the coding of each audio sample. A vocalization was defined as the production of an audible vocal sound by one of the partners where interruption of sound did not exceed 300 ms (criteria from Gratier et al., 2015). Overlapping vocalizations were defined as the production of audible and partially simultaneous vocalizations of the two partners. Maternal pauses were identified when maternal vocalizations were interrupted for longer than 300 ms. All maternal vocalizations and pauses were coded and quantified in terms of frequency (number of vocalizations per minute), duration (in ms) and coefficients of variation were obtained in both singing and speaking conditions. All infants' vocalizations were coded and quantified in terms of frequency of overlapping vocalizations, of vocalizations occurring during maternal vocal pauses and of total vocalizations (sum of overlapping vocalizations and vocalizations in pauses). Infants' vocalizations during pauses were considered to be temporally contingent on a maternal vocalization if the duration between the end of a maternal vocalization and the beginning of the infant's vocalization was not greater than 3000 ms (Gratier et al., 2015). This was taken as the maximum time window for social contingency in both the speaking and singing conditions. Infants' vocalizations were coded for all periods (baseline, speaking/singing, offsets). Latencies of infants' vocalizations in relation to the end of a maternal vocalization were measured (in ms).

### 2.7. Reliability

An independent coding of the data was performed by two researchers for 30% of the mothers' and babies' vocal variables. Inter observer reliability was estimated *via* intraclass correlation coefficients (ICC). For frequencies of vocalizations and pauses in maternal speaking and singing high agreement coefficients were found: 1) speech vocalizations, ICC = .991,  $p = .000$ ; 2) pauses in speech, ICC = .985,  $p = .000$ ; 3) singing vocalizations, ICC = .999,  $p = .000$  and 4) pauses in singing, ICC = .999,  $p = .000$ . Concerning babies' vocalizations, coefficients were also high: a) vocalizations during speaking pauses, ICC = .982,  $p = .000$ , b) overlapping during speaking, ICC = .985,  $p = .000$ , c) vocalizations during pauses in singing, ICC = .988,  $p = .000$  and overlapping in singing, ICC = .996,  $p = .000$ .

### 2.8. Statistical analysis

A non-parametric test (Mann Whitney) was used to observe if there were differences between maternal vocalizations depending on the order, singing-speaking or speaking-singing. The same was used to compare babies' vocalizations in the two groups. Because differences were not significant it was decided to combine the data independently of the order of the conditions. A *T*-test for paired samples was performed to compare maternal vocal variables in speaking and singing conditions once it was established that nearly all maternal variables followed the normal distribution (Kolmogorov-Smirnov Test). Because babies' vocalizations were not close to

the normal distribution, a nonparametric test (Wilcoxon) for paired samples was used to compare babies' vocalizations according to the different test periods: a) baseline vs. all other periods, b) speaking vs. singing, c) speaking vs. offset speaking, d) singing vs. offset singing and e) offset speaking vs. offset singing. A multiple linear regression was performed to observe the prediction of babies' vocalizations during baseline, singing and speaking conditions as well as the prediction of maternal vocalizations in the same conditions.

### 3. Results

#### 3.1. Overall vocal production

For the entire sample ( $N = 36$ ) a total of 3953 maternal vocalizations was identified of which 2186 were vocalizations produced in the 3-min speaking condition and 1767 were vocalizations produced in the 3-min singing condition. In total 3900 pauses were identified of which 2160 were pauses during the speaking condition and 1740 were pauses during the singing condition.

In Table 1 the descriptive statistics of infants' vocalizations during protocol are displayed. Overall ( $N = 36$ ), a total of 1172 infants' vocalizations was identified among which 306 were vocalizations during the baseline periods. A total of 205 vocalizations were recorded in the speaking condition and 189 vocalizations were recorded in the singing condition. In total, 258 vocalizations were identified during the speaking offset and 214 vocalizations were identified during the singing offset. Within the sample of infants' vocalizations a total of 228 overlapping vocalizations was identified of which 86 occurred during maternal speaking and 142 during maternal singing. Furthermore, 166 infants' vocalizations were identified as occurring during pauses among which 119 in the speaking condition and 47 in the singing condition.

#### 3.2. Temporal characteristics of maternal vocalizations and pauses in the speaking and singing conditions

A series of *t*-Test analysis was performed to compare frequencies, durations and variability of the durations of maternal vocalizations and pauses in the singing and speaking conditions. A Bonferroni Correction was used ( $.01/6 = .002$ ) and all the comparisons remained significant. As shown in Table 2, we observed a significantly higher frequency of maternal vocalizations in the speaking condition compared to the singing condition [ $t(35) = 5.776$ ,  $p = .000$ ]. Accordingly, a higher frequency of pauses was observed in the speaking relatively to the singing condition [ $t(35) = 5.799$ ,  $p = .000$ ]. The mean length of maternal vocalizations was higher in singing than in speaking [ $t(35) = -10.465$ ,  $p = .000$ ]. However, the mean length of maternal pauses between vocalizations was significantly lower in singing than in speaking [ $t(35) = 8.153$ ,  $p = .000$ ]. The coefficient of variation of length of maternal vocalizations was significantly higher in speaking than in singing [ $t(35) = 8.768$ ,  $p = .000$ ]. Also the coefficient of variation is significantly higher for length of pauses during speaking than during singing [ $t(34) = 5.888$ ,  $p = .000$ ]. It should be noted that the mean coefficient of variation of length of maternal vocalizations is greater than 30% in both speech and in singing pointing to a generalized variability. However, when compared to speaking, maternal singing to babies in the NICU is characterized by fewer vocalizations and pauses, longer vocalizations, shorter pauses and greater regularity of vocalizations and pauses duration.

**Table 1**  
Frequencies of infants' vocalizations according to maternal vocal condition.

	Baseline	Speaking	Speaking offset	Singing	Singing offset	Overlap in speaking	Overlap in singing	Pauses in speaking	Pauses in singing
M	8.50	5.66	7.11	5.61	5.89	2.39	4.30	3.03	1.30
SD	9.91	6.19	10.76	6.64	6.95	3.35	5.33	3.24	1.90
min.-max.	0-43	0-24	0-51	0-26	0-32	0-15	0-21	0-14	0-7

**Table 2**  
Analysis of differences of maternal vocalizations and maternal pauses between speaking and singing conditions for total sample ( $N = 36$ ).

Vocalizations and pauses	total sample ( $N = 36$ )								
	frequency			length			coefficient of variation (%)		
	M	SD	min.-max.	M	SD	min.-max.	M	SD	min.-max.
MSV	60.72	10.37	44-83	1.33	.323	.54-2.39	59.77	10.26	42.91-88.22
MSGV	49.08	11.33	20-72	2.84	.774	1.77-4.84	34.46	14.95	7.64-74.18
MSV vs. MSGV	$t = 5.776$ $p = .000^*$			$t = -10.465$ $p = .000^*$			$t = 8.768$ $p = .000^*$		
MSP	60.00	10.35	44-82	1.66	.515	.78-2.63	54.88	10.09	29.20-117
MSGP	48.33	11.34	28-71	.979	.293	.57-1.88	32.15	11.95	16.72-77.51
MSP vs. MSGP	$t = 5.799$ $p = .000^*$			$t = 8.153$ $p = .000^*$			$t = 5.888$ $p = .000^*$		

MSV – maternal speaking vocalizations; MSGV – maternal singing vocalizations; MSP – maternal speaking pauses; MSGP – maternal singing pauses. \* Bonferroni Correction ( $.01 / 6 = .002$ ).

### 3.3. Characteristics of infants' vocalizations according to maternal vocal condition

Descriptive results about infants' vocalizations according to maternal vocal condition are presented in Table 1. A nonparametric test (Wilcoxon) was used to compare infants' vocalizations according to maternal vocal condition. Results are presented in Table 3.

Compared to the baseline condition (baseline:  $M = 8.5$ ,  $SD = 9.91$ ), infants vocalized significantly less frequently in the singing condition ( $Z = -2.032$ ,  $p = .042$ ; singing:  $M = 5.61$ ,  $SD = 6.64$ ) and in the speaking condition ( $Z = -1.966$ ,  $p = .049$ ; speaking:  $M = 5.66$ ,  $SD = 6.19$ ). Also when comparing to baseline, they vocalized less frequently during the singing offset ( $Z = -2.102$ ,  $p = .036$ ; singing offset:  $M = 5.89$ ,  $SD = 6.95$ ) but no significant difference was found when comparing with speaking offset. We did not find a significant difference in infant's vocalizations frequency between the singing and speaking conditions. However, we found a positive correlation between frequencies of vocalizations in the singing and in the speaking conditions ( $r = .368$ ,  $p = .027$ ). We also found that overlapping vocalizations in the singing condition ( $M = 4.30$ ,  $SD = 5.33$ ) are significantly higher ( $Z = -2.688$ ,  $p = .007$ ) than overlapping vocalizations in the speaking condition ( $M = 2.39$ ,  $SD = 3.35$ ). In addition, infants' vocalizations occurred more frequently during pauses in speaking ( $M = 3.03$ ,  $SD = 3.24$ ) than during pauses in singing ( $M = 1.30$ ,  $SD = 1.90$ ) and this difference is significant ( $Z = -2.912$ ,  $p = .004$ ). There was a significantly higher frequency ( $Z = -2.688$ ,  $p = .007$ ) of overlapping vocalizations in singing ( $M = 4.30$ ,  $SD = 5.33$ ) than in speaking ( $M = 2.39$ ,  $SD = 3.35$ ). In a global view, preterm infants vocalize less often while their mothers speak and sing than during baseline and their vocalizations tend to be more alternating in the speaking condition and more overlapping in the singing condition. Maternal singing prompts the baby to overlap more while maternal speaking seems to induce more vocalizations in pauses.

A further analysis was carried out on the latency times (ms) of infants' vocalizations occurring during pauses (between 300 ms and 3000 ms) in speaking and singing. Results show that the average latency times are significantly higher ( $Z = -2.373$ ,  $p = .018$ ) in speaking ( $M = .8662$  ms,  $SD = .636$ ,  $min. = .04$ ,  $max. = 3$ ) than in singing ( $M = .5454$  ms,  $SD = .366$ ,  $min. = .04$ ,  $max. = 1.54$ ). It is concluded that preterm infants take more time to respond to speech utterances than to sung nonverbal vocalizations.

### 3.4. Prediction of frequency of infants' vocalizations in speaking and singing conditions

To know which variables best predict babies' vocalizations in the baseline, a multiple linear regression (stepwise) was performed. Results show that among sociodemographic and clinical variables mothers' education is the only good predictor of frequency of infants' vocalizations in baseline ( $F$  change = 4.990,  $p = .033$ ). Also a significant negative correlation was found between frequency of infants' vocalizations in the speaking condition and maternal education ( $r = -.409$ ,  $p = .05$ ). Furthermore, frequency of infants' vocalizations during speaking pauses and mothers' education were also negatively correlated ( $r = -.419$ ,  $p = .05$ ).

A multiple hierarchic linear regression (enter) analysis was performed in order to test if some specific independent variables are or not good predictors of infants' vocalizations. We supposed that during singing condition the infants' vocalizations are influenced by the frequency of maternal vocalizations. In a univariate linear regression this was confirmed ( $R^2 = .129$ ,  $F$  change = 5.036,  $p = .031$ ). Also the infants' overlapping vocalizations in singing condition were explained by the frequency of maternal vocalizations ( $R^2 = .131$ ,  $F$  change = 5.109,  $p = .030$ ). It is also important to control the possible effects of other variables (sociode-

**Table 3**

Nonparametric tests of differences in infants' vocalizations frequencies according to maternal vocal condition.

	Baseline	Speaking	Speaking offset	Singing	Singing offset	Overlap in speaking	Overlap in singing	During pauses in speaking	During pauses in singing
Speaking	-1.966 ( $p = .049$ )*								
Speaking offset	-1.532 ( $p = .126$ )	-.602 ( $p = .547$ )			-.216 ( $p = .829$ )				
Singing	-2.032 ( $p = .042$ )*	-.131 ( $p = .896$ )							
Singing offset	-2.102 ( $p = .036$ )*			-.628 ( $p = .530$ )					
Overlap in speaking									
Overlap in singing						-2.688 ( $p = .007$ )**			
During pauses in speaking									
During pauses in singing								-2.912 ( $p = .004$ )**	

\*  $p < .05$ .

\*\*  $p < .01$ .

mographic and clinical). In this sense, mothers' age and education, number of fetuses during the last gestation, baby's chronological age at observation, the sequence "singing-speech" or "speech-singing" and baby's gender were submitted to univariate linear regressions. Among the sociodemographic and clinical variables only the baby's gender was able to explain the infants' vocalizations in singing condition ( $R^2 = .162$ ,  $F$  change = 6.59,  $p = .015$ ). So, in the multiple hierarchic linear regression (enter) analysis, as dependent variables we introduced, one for each regression, all variables concerning infants' vocalizations in the speaking and singing conditions: 1) overall frequency of vocalizations, 2) frequency of overlapping vocalizations and 3) vocalizations occurring during pauses (lasting 300–3000 ms). As independent variables, the already mentioned sociodemographic and clinical variables were controlled and introduced as predictor variables: Model 1) mothers' age and education, Model 2) number of fetuses in last pregnancy and infants' chronological age at observation (number of days since birth at observation), Model 3) infants' gender, Model 4) order of conditions and Model 5) mothers' vocal variables (one for each analysis). As mothers' vocal variables we considered frequencies, lengths and coefficients of variation of lengths in speaking, in singing, maternal pauses in speaking and maternal pauses in singing. Results are presented in Table 4.

As shown in Table 4, two predictor variables yielded significant results for the explanation of two dependent variables concerning infants' vocalizations during singing. Frequency of maternal vocalizations in singing was a good predictor of infants' vocalizations in singing ( $p = .015$ ) as well as of overlapping vocalizations during singing ( $p = .000$ ). Frequency of maternal pauses in singing was also a significant predictor of infants' vocalizations in singing ( $p = .016$ ) and of overlapping vocalizations in singing ( $p = .011$ ).

Among the sociodemographic and clinical variables, infants' gender was a good predictor of infants' vocalizations in the singing condition. Nonparametric analysis for independent samples (Mann Whitney) was performed to compare infants' vocalizations in the singing condition according to infants' gender, but the results were not significant. However there was a negative and significant correlation ( $r = -.403$ ,  $p = .015$ ) between infants' gender and infants' vocalizations in singing. It was concluded that girls present a higher frequency of vocalizations in the singing condition when compared to boys. Mothers' education is another significant predictor of all infants' variables in the speaking condition, namely: frequency of infants' vocalizations ( $p = .013$ ), frequency of overlapping vocalizations ( $p = .042$ ) and frequency of vocalizations during speech pauses ( $p = .016$ ). A negative and significant correlation ( $r = -.409$ ,  $p = .05$ ) between the frequency of infants' vocalizations in the speaking condition and mothers' education (number of successful years at school) was found. In much the same way, another negative and significant correlation was found between infants' vocalizations in speaking pauses and mothers' education ( $r = -.419$ ,  $p = .05$ ).

Among vocal maternal variables, frequency of maternal vocalizations in singing is the only predictor of babies' vocalizations in singing condition. Significant positive correlations were found in the singing condition between: a) frequency of maternal and infants' vocalizations ( $r = .359$ ,  $p = .031$ ), b) frequency of maternal vocalizations and overlapping vocalizations ( $r = .360$ ,  $p = .031$ ), c) maternal singing pauses and infant vocalizations ( $r = .354$ ,  $p = .034$ ) and d) maternal singing pauses and overlapping vocalizations ( $r = .356$ ,  $p = .033$ ).

**Table 4**  
Multiple linear regression analysis, DV-babies' vocal variables, IV-maternal vocal variables.

Model	R	R <sup>2</sup>	Adjusted R <sup>2</sup>	St. error of the estimate	R <sup>2</sup> change	F change	df1	df2	Sig. of F change
babies' vocalizations in singing (DV) and frequency of maternal vocalizations in singing (IV)									
Model 1	.222	.049	-.008	6.67043	.049	.857	2	33	.434
Model 2	.259	.067	-.053	6.81820	.018	.293	2	31	.748
Model 3	.489	.239	.112	6.25865	.172	6.791	1	30	.014
Model 4	.520	.271	.120	6.23195	.032	1.258	1	29	.271
Model 5*	.642	.412	.265	5.69589	.141	6.715	1	28	.015
babies' vocalizations in singing (DV) and frequency of maternal pauses in singing (IV)									
Model 1	.222	.049	-.008	6.67043	.049	.857	2	33	.434
Model 2	.259	.067	-.053	6.81820	.018	.293	2	31	.748
Model 3	.489	.239	.112	6.25865	.172	6.791	1	30	.014
Model 4	.520	.271	.120	6.23195	.032	1.258	1	29	.271
Model 5**	.641	.410	.263	5.70316	.140	6.627	1	28	.016
babies' overlapping vocalizations in singing (DV) and frequency of maternal vocalizations in singing (IV)									
Model 1	.200	.040	-.018	5.38207	.040	.686	2	33	.511
Model 2	.224	.050	-.072	5.52275	.010	.170	2	31	.844
Model 3	.447	.200	.067	5.15274	.150	5.612	1	30	.024
Model 4	.514	.264	.112	5.02659	.064	2.525	1	29	.123
Model 5*	.980	.960	.951	1.18559	.696	493.287	1	28	.000
babies' overlapping vocalizations in singing (DV) and frequency of maternal pauses in singing (IV)									
Model 1	.200	.040	-.018	5.38207	.040	.686	2	33	.511
Model 2	.224	.050	-.072	5.52275	.010	.170	2	31	.844
Model 3	.447	.200	.067	5.15274	.150	5.612	1	30	.024
Model 4	.514	.264	.112	5.02659	.064	2.525	1	29	.123
Model 5**	.648	.420	.275	4.54260	.156	7.509	1	28	.011

DV - dependent variables; IV - independent variables; Model 1: mothers' age, education, Model 2: number of fetus in last pregnancy, baby's chronological age at observation, Model 3: babies' gender, Model 4: protocol sequence, Model 5\*: frequency of maternal vocalizations in singing; Model 5\*\*: frequency of maternal pauses in singing.

### 3.5. Prediction of frequency of maternal vocalizations in speaking and singing conditions

Finally, a multiple hierarchic linear regression (enter) was also performed to identify which of the infants' vocal, sociodemographic and clinical variables were good predictors of maternal vocalizations frequencies in speaking and singing conditions. A preliminary univariate analysis of linear regression was performed. It was observed that infants' overlapping vocalizations in singing constitute a good predictor variable of the frequency of maternal vocalizations in singing ( $R^2 = .131$ ,  $F$  change = 5.109,  $p = .030$ ). About sociodemographic and clinical variables only the infants' chronological age at observation was a good predictor of the frequency of maternal vocalizations in singing ( $R^2 = .119$ ,  $F$  change = 4.582,  $p = .040$ ).

For the multiple hierarchic linear regression (enter) the first four models were exactly the models used in 3.4. Results are displayed in Table 5. The frequency of infants' overlapping vocalizations in singing is a good predictor ( $p = .009$ ) of the frequency of maternal vocalizations in singing. Among sociodemographic and clinical variables, infants' chronological age at observation was the only variable able to contribute to the explanation of the dependent variable ( $p = .049$ ). According to results, none of sociodemographic variables, clinical variables and variables about babies' vocalizations in speaking condition was a significant predictor of maternal vocalizations in speaking. A significant positive correlation was found, in the singing condition, between the frequency of overlapping vocalizations and the frequency of maternal vocalizations ( $r = .360$ ,  $p = .031$ ) while a significant negative correlation was found between infant's chronological age at observation and frequency of maternal vocalizations in singing ( $r = -.345$ ,  $p = .040$ ).

## 4. Discussion and conclusion

This study had two main aims. First, to describe the temporal organization of maternal speech and singing addressed to preterm newborns during skin-to-skin kangaroo care and, second, to determine the effects of maternal addressed vocal communication on infants' vocalization rate. Our study reveals a clear difference in the temporal characteristics of maternal infant-directed speech and singing. When compared to speech, maternal singing was characterized by fewer vocalizations and pauses, and by longer length of vocalizations and shorter length of pauses. Despite the variability in the length of both speech and singing vocalizations we observed a greater regularity of vocalization and pause lengths in singing than in speech.

With regard to the rates of infants' vocalizations, our results showed that both maternal singing and speaking conditions induced a decrease in infants' vocalizations compared to a silent baseline. Our findings suggest that the two modes of vocal expression mothers can use to communicate with their preterm infants are linked to different temporal frameworks based on a variable parsing of sound and silence. Infant-directed speaking and singing may serve different purposes for preterm infants. It is known that speaking and singing to full-term and older infants elicit different responses, with infants showing greater sustained attention to infant-directed singing than to infant-directed speaking (Nakata & Trehub, 2004). Yet it was also found that preterm infants exposed to live infant-directed speaking and singing by their mothers in the NICU spend more time in an quiet alert state when their mothers speak to them and in a sleep state when their mothers sing to them (Filippa et al., 2013). In our study the decrease in the infants' vocalizations rate during maternal vocal communication in both conditions may be interpreted as the result of the enhanced attention on the part of the infants. However, a decrease in the vocalizations rate during the singing offset relatively to the baseline was also found. It is possible that the experience of being sung do induced a change in the infants' behavioral state, leading to a higher proportion of drowsiness or sleep indexed by reduced vocalization.

In a singing mode, durations of phrases are constrained by musical structure and are thus less variable than in the speaking mode when vocalizations' durations involve an active chunking of syllables by the speaker. However, in this study mothers exclusively sang improvised tunes and did not use words in their singing. Thus the singing condition in this study also involved active chunking of hummed syllables and vowels. It is interesting, in this context, that singing involved less variability in vocal sound duration than speaking. It is possible that in the singing mode mothers produced phrase-like sequences of sound, reproducing the musical structure of traditional baby songs. The variability of infant-directed speech has been described as expressive timing, that is, a well-balanced irregularity that favors infant attention and participation (Gratier, 2003).

Based on the difference in expressive timing we have just reported, and because we found that pauses were generally shorter in the infant-directed singing mode than in infant-directed speaking, we expected infants to vocalize more often during infant-directed

**Table 5**

Multiple hierarchic linear regression for dependent variables concerning maternal vocalizations in singing (DV) and babies' overlapping vocalizations in singing (IV).

Model: predictors	R	R <sup>2</sup>	Adjusted R <sup>2</sup>	St. error of the estimate	R <sup>2</sup> change	F change	df1	df2	Sig. of F change
Model 1	.318	.101	.047	11.06186	.101	1.858	2	33	.172
Model 2	.419	.176	.069	10.93077	.074	1.398	2	31	.262
Model 3	.421	.177	.040	11.10296	.001	.046	1	30	.832
Model 4	.450	.203	.038	11.11512	.026	.934	1	29	.342
Model 5	.614	.377	.221	10.00009	.174	7.828	1	28	.009

DV - dependent variables; IV - independent variable; Model 1: mothers' age, education, Model 2: number of fetus in last pregnancy, baby's chronological age at observation, Model 3: babies' gender, Model 4: protocol sequence, Model 5: frequency of babies' overlapping vocalizations in singing.

speaking. Although, we did find that infants vocalized less frequently in both vocal communication conditions than in the baseline condition, we did not find any difference in vocalization rate between the speaking and singing conditions. However, infants were found to vocalize more often during pauses in the speaking condition than during pauses in the singing condition. When comparing latency times (occurred during pauses) of babies' vocalizations in both conditions, we found that infants were slower to respond to a speech vocalization than to a hummed vocalization. These differences in temporal organization of maternal speech and singing seem to induce a variable vocal responsiveness in preterm babies. Thus it may be that preterm infants do differentiate between speech and singing; if so, this early competence may be based on features such as temporal organization and phonological structure. However, infants' vocalizations may not be an adequate measure of infants' responsiveness to maternal vocal expression as it requires much effort on the part of infants and thus involves important inter and intra-individual variability. Physiological measures also registered at the same sample may be helpful to elucidate this question in a near future.

Results showed that the singing condition is associated with a higher frequency of overlapping vocalizations whilst the speaking condition induces a higher frequency of infants' vocalizations occurring during pauses. Results also showed a correlation between overlapping vocalizations during maternal singing and the number of vocalizations and pauses in maternal singing. This suggests that preterm babies are sensitive to variations in temporal units of maternal singing; the more the maternal singing is segmented the more babies' overlapping vocalizations will be observed. We think that these observations should be discussed according to the behavioral state of the preterm baby. This conclusion should be taken into account while parents are being encouraged to sing in a contingent way to their preterm babies.

A recent study on newborns' vocal production during addressed maternal speech suggests that a 1000 ms temporal window is the most fitting criteria to appreciate social contingency in the neonatal period (Dominguez et al., 2016). Our results showed that latency times of preterm infants' contingent vocal responses in the speaking condition are significantly different from latency times observed in the singing condition. In the speaking condition responses occur within a latency time close to 1000 ms, as in the Dominguez et al. (2016) study, while in the singing condition responses occur with an average latency time of 500 ms. Preterm babies thus appear to take more time to respond vocally to speech than to hummed tunes, suggesting that singing affords a tighter synchrony between infants and mothers. As in singing condition it is also observed a higher frequency of overlapping vocalizations, it is possible that a higher co-modulation in mother-baby vocal exchanges during singing favors an earlier attunement in preterm dyads, in the sense of affect attunement (Stern, 1985).

Our results also show that infants' gender is the best predictor of infants' vocalizations rate during maternal singing. It was observed that female babies are prone to vocalize more frequently than male babies during the singing condition. Mothers' education is also an important predictor of infants' vocalizations. Results show that there is an increase of vocalizations rate during speech and also in the baseline condition for infants of mothers with fewer years of education. Temporal and prosodic features of these mothers' speech would be likely candidates to explain these findings. However our results show that no variable of maternal speech is a good predictor of infants' vocalizations rate in the speaking condition. Among maternal vocal variables, rate of maternal vocalizations in singing is the only significant predictor of infants' vocalizations rate in the singing condition. The more mothers segment their singing the more often infants vocalize during singing. Surprisingly, vocalizations' length, variability of vocalizations' lengths as well as the length of pauses and its variability in maternal singing were not predictive of infants' vocalizations rate during singing. Results suggest that infants are more sensitive to the segmentation of singing than to the length of phrases or to their expressive timing. Results also showed that infants' overlapping vocalizations in the singing condition predicts maternal vocalizations rate in singing; the more babies overlap during maternal singing the more their mothers segment their vocalizations in singing. Possibly this is in line with data about co-modulation happening during mother-infant vocal exchanges in preterm dyads (Filippa et al., 2013).

Infants' chronological age is another important predictor of specific temporal features in maternal singing; the younger the infant is the more the mother segments her singing, leading to an increase of maternal vocalization rate. It was concluded that mothers adapt the flow of their singing to infants' age and that infants induce an increase in maternal singing segmentation by producing more overlapping vocalizations. This suggests that the singing condition affords a form of engagement in preterm babies paving the way for a co-modulation during vocal exchanges. If this is so, maternal singing might play an important role in bonding for preterm dyads, which was already observed in term dyads (Fancourt & Perkins, 2018; Persico et al., 2017).

Overall, this study shows that preterm infants are highly responsive to their mothers' elicitation through close physical contact and vocal expression. No study to date has shown that preterm infants vocalize contingently and differentially according to the mode of vocal communication addressed to them, speaking or singing. It is important to explore the extent to which different modes of vocal expression induce states of interpersonal synchrony and how they affect states of attention and awareness in preterm infants. In a near future we will present results including physiological measures of both infants and mothers as well as about the infants' behavioral states in both conditions observed with the dyads of this sample.

#### Authors' contribution

The first five authors contributed to the design of the research, the fourth author validated the inclusion and exclusion criteria and the fifth author contributed to the organization of the setting namely in kangaroo care. The first author collected and coded the data, the first two authors performed the statistical analysis and wrote the text of the paper. The third and the last authors reviewed the text and contributed to wrote the final version of the paper.

## Uncited references

Bertoncini et al. (1995), Stern (1999) and Stern et al. (1985).

## Declaration of Competing Interest

The authors declare no conflict of interest, financial or otherwise.

## Acknowledgements

The reliability of the coding process was supported by Carolina Sá a student of the Master of Musicology and a cooperative member of CESEM-NOVA-FCSH. The first author is currently supported by a Post-PhD. scholarship of the Strategic Project of CESEM (UID/EAT/ 00693/2013) financed by FCT (Portuguese Foundation for Science and Technology). The paper was submitted to linguistic review by a translator financed by CESEM. The authors would like to thank to the Ethical Committee of the Central Lisbon Hospital Center as well as to the nursing team that actively engaged in the process.

## References

- Arnon, S., Diamant, C., Bauer, S., Regev, R., Sirota, G., Litmanovitz, I., 2014. Maternal singing during kangaroo care led to autonomic stability in preterm infants and reduced maternal anxiety. *Acta Paediatrica* 103, 1039–1044. <https://doi.org/10.1111/apa.12744>.
- Bertoncini, J., Floccia, C., Nazzi, T., Mehler, J., 1995. Morae and syllables: Rhythmical basis of speech representations in neonates. *Language and Speech* 38 (4), 311–329. <https://doi.org/10.1177/002383099503800401>.
- Best, K., Bogossian, F., New, K., 2018. Language exposure of preterm infants in the neonatal unit: A systematic review. *Neonatology* 114, 261–276. <https://doi.org/10.1159/000489600>.
- Bateson, M.C., 1979. The epigenesis of conversational interaction: A personal account of research development. In: Bullowa, M. (Ed.), *Before speech: The Beginning of human communication*. Cambridge University Press, London, pp. 63–77.
- Brazelton, T.B., Nugent, J.K., 2011. *The neonatal behavioral assessment scale*, 4th edition Mac Keith Press, London.
- Bruner, J.S., 1983. *Child's talk. Learning to use language*. Norton, New York.
- Buil, A., Carchon, I., Apter, G., Laborne, F.X., Gratier, M., Devouche, E., 2016. Kangaroo supported diagonal flexion positioning: New insights into skin-to-skin contact for communication between mothers and very preterm infants. *Archives de Pédiatrie* 23, 913–920. <https://doi.org/10.1016/j.arcped.2016.04.023>.
- Byers-Heinlein, K., Burns, T.C., Werker, J.F., 2010. The roots of bilingualism in newborns. *Psychological Science* 21 (3), 343–348. <https://doi.org/10.1177/0956797609360758>.
- Caskey, M., Stephens, B., Tucker, R., Vohr, B., 2011. Importance of parent talk on the development of preterm infant vocalizations. *Pediatrics* 128 (5), 910–916. <https://doi.org/10.1542/peds.2011-0609>.
- Caskey, M., Stephens, B., Tucker, R., Vohr, B., 2014. Adult talk in the NICU with preterm infants and developmental outcomes. *Pediatrics* 133, e578–e584.
- DeCasper, A.J., Fifer, W.P., 1980. Of human bonding: Newborns prefer their mother's voices. *Science* 208, 1174–1176. <https://doi.org/10.1126/science.7375928>.
- DeCasper, A.J., Spence, M.J., 1986. Prenatal maternal speech influences newborns' perception of speech sound. *Infant Behavior and Development* 9, 133–150. [https://doi.org/10.1016/0163-6383\(86\)90025-1](https://doi.org/10.1016/0163-6383(86)90025-1).
- DeCasper, A.J., Lecanuet, J.-P., Busnel, M.-C., Granier-Deferre, C., Maugeais, R., 1994. Fetal reactions to recurrent maternal speech. *Infant Behavior and Development* 17, 159–164. [https://doi.org/10.1016/0163-6383\(94\)90051-5](https://doi.org/10.1016/0163-6383(94)90051-5).
- Delavenne, A., Gratier, M., Devouche, E., 2013. Expressive timing in infant-directed singing between 3 and 6 months. *Infant Behavior and Development* 36 (1), 1–13. <https://doi.org/10.1016/j.infbeh.2012.10.004>.
- Dominguez, S., Devouche, E., Apter, G., Gratier, M., 2016. The roots of turn-taking in the neonatal period. *Infant and Child Development* 25, 240–255. <https://doi.org/10.1002/icd.1976>.
- Falk, S., Kello, C., 2017. Hierarchical organization in the temporal structure of infant-direct speech and song. *Cognition* 163, 80–86. <https://doi.org/10.1016/j.cognition.2017.02.017>.
- Fancourt, D., Perkins, R., 2018. The effects of mother–infant singing on emotional closeness, affect, anxiety, and stress hormones. *Music & Science* 1, 1–10. <https://doi.org/10.1177/2059204317745746>.
- Ferrari, G.A., Nicolini, Y.D., Tosato, E., Hussain, C., Scesa, M., Romei, E., et al., 2016. Ultrasonographic investigation of human fetus responses to maternal communicative and non-communicative stimuli. *Frontiers in Psychology* 3 (16), 1–10. <https://doi.org/10.3389/fpsyg.2016.00354>.
- Filippa, M., Devouche, E., Arioni, C., Imberty, M., Gratier, M., 2013. Live maternal speech and singing have beneficial effects on hospitalized preterm infants. *Acta Paediatrica* 102, 1017–1020. <https://doi.org/10.1111/apa.12356>.
- Filippa, M., Panza, C., Ferrari, F., Frassoldati, R., Kuhn, P., Balduzzi, S., et al., 2017. Systematic review of maternal voice interventions demonstrates increased stability in preterm infants. *Acta Paediatrica* 106 (8), 1220–1229. <https://doi.org/10.1111/apa.13832>.
- Ginsburg, G.P., Kilbourne, B.K., 1988. Emergence of vocal alternation in mother-infant interchanges. *Journal of Child Language* 15, 221–235. <https://doi.org/10.1017/S0305000900012344>.
- Goldstein, M.H., Schwade, J.A., 2008. Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science* 19 (5), 515–523.
- Granier-Deferre, C., Ribeiro, A., Jacquet, A.-Y., Bassereau, S., 2011. Near-term fetuses process temporal features of speech. *Developmental Science* 14 (2), 336–352. <https://doi.org/10.1111/j.1467-7687.2010.00978.x>.
- Gratier, M., 2003. Expressive timing and interactional synchrony between mothers and infants: Cultural similarities, cultural differences, and the immigration experience. *Cognitive Development* 18, 533–554. <https://doi.org/10.1016/j.cogdev.2003.09.009>.
- Gratier, G., Devouche, E., Guellai, B., Infanti, R., Yilmaz, E., Parlato-Oliveira, E., 2015. Early development of turn-taking in vocal interaction between mothers and infants. *Frontiers in Psychology* 6, 1167. <https://doi.org/10.3389/fpsyg.2015.01167>.
- Gratier, M., Trevarthen, C., 2008. Musical narrative and motives for culture in mother-infant vocal interaction. *The Journal of Consciousness Studies* 15 (10–11), 122–158.
- Hilbrink, E.E., Gattis, M., Levinson, S.C., 2015. Early developmental changes in the timing of turn-taking: A longitudinal study of mother-infant interaction. *Frontiers in Psychology* 6, 1492. <https://doi.org/10.3389/fpsyg.2015.01492>.
- Jaffe, J., Beebe, B., Feldstein, S., Crown, C.L., Jasnaw, M.D., 2001. Rhythms of dialogue in infancy: Coordinated timing in development. *Monographs of the Society for Research in Child Development* 66, 1–8. <https://doi.org/10.1111/1540-5834.00137>.
- Jardri, R., Houffine-Debarge, V., Delion, P., Pruvo, J.-P., Thomas, P., Pins, D., 2012. Assessing fetal response to maternal speech using a noninvasive functional brain imaging technique. *International Journal of Developmental Neuroscience* 30, 159–161. <https://doi.org/10.1016/j.ijdevneu.2011.11.002>.

- Keller, H., Lohaus, A., Völker, S., Cappenberg, M., Chasiotis, A., 1999. Temporal contingency as an independent component of parenting behavior. *Child Development* 70, 474–485. <https://doi.org/10.1111/1467-8624.00034>.
- Kisilevsky, B., Hains, S., 2011. Onset and maturation of fetal heart rate response to the mother's voice over late gestation. *Developmental Science* 14 (2), 214–223. <https://doi.org/10.1111/j.1467-7687.2010.00970.x>.
- Lavelli, M., Fogel, A., 2013. Interdyad differences in early mother–infant face-to-face communication: Real-time dynamics and developmental pathways. *Developmental Psychology* 49, 2257–2271. <https://doi.org/10.1037/a0032268>.
- Lee, G.Y., Kisilevsky, B.S., 2014. Fetuses respond to father's voice but prefer mother's voice after birth. *Developmental Psychobiology* 56 (1), 1–11. <https://doi.org/10.1002/dev.21084>.
- Levinson, S.C., 2016. Turn-taking in human communication—Origins and implications for language processing. *Trends in Cognitive Sciences* 20 (1), 6–14. <https://doi.org/10.1016/j.tics.2015.10.010>.
- Malloch, S., Trevarthen, C. (Eds.), 2009. *Communicative musicality: Exploring the basis of human companionship*. Oxford University Press, Oxford.
- Maturana, H.R., 1978. Biology of language: The epistemology of reality. In: Miller, G.A., Lenneberg, E. (Eds.), *Psychology and biology of language and thought: Essays in honor of Eric Lenneberg*. Academic Press, New York, pp. 27–63.
- May, L., Byers-Heinlein, K., 2011. Language and the newborn brain: Does prenatal language experience shape the neonate neural response to speech?. *Frontiers in Psychology* 2, 222. <https://doi.org/10.3389/fpsyg.2011.00222>.
- Moon, C., Lagercrantz, H., Kuhl, P.K., 2013. Language experienced in utero affects vowel perception afterbirth: A two-country study. *Acta Paediatrica* 102, 156–160. <https://doi.org/10.1111/apa.12098>.
- Nakata, T., Trehub, S.E., 2004. Infants' responsiveness to maternal speech and singing. *Infant Behavior and Development* 27, 455–464. <https://doi.org/10.1016/j.infbeh.2004.03.002>.
- Oller, D.K., 2000. *The emergence of the speech capacity*. Psychology Press, Hove.
- Marx, V., Nagy, E., 2015. Fetal behavioural responses to maternal voice and touch. *PLoS One* 10 (6), e0129118. <https://doi.org/10.1371/journal.pone.0129118>.
- Papousek, H., Papousek, M., Bornstein, M.H., 1985. The naturalistic environment of young infants: On the significance of homogeneity and variability in parental speech. In: Field, T., Fox, N. (Eds.), *Social perception in infants*. Ablex Publishing Corporation, New York, pp. 269–297.
- Pelaez, M., Virués-Ortega, J., Gewirtz, J.L., 2011. Contingent and noncontingent reinforcement with maternal vocal imitation and motherese speech: Effects on infant vocalizations. *European Journal of Behavior Analysis* 12, 277–287. <https://doi.org/10.1080/15021149.2011.11434370>.
- Persico, G., Antolini, L., Vergani, P., Costantini, W., Nardi, M.T., Bellotti, L., 2017. Maternal singing of lullabies during pregnancy and after birth: Effects on mother–infant bonding and on newborns' behaviour. *Concurrent Cohort Study. Women and Birth* 30, e214–e220. <https://doi.org/10.1016/j.wombi.2017.01.007>.
- Ramus, F., Nespor, M., Mehler, J., 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 73, 265–292. [https://doi.org/10.1016/S0010-0277\(99\)00058-X](https://doi.org/10.1016/S0010-0277(99)00058-X).
- Reigado, J., Rocha, A., Rodrigues, H., 2011. Vocalizations of infants (9–11 months old) in response to musical and linguistic stimuli. *International Journal of Music Education* 29 (3), 241–256. <https://doi.org/10.1177/0255761411408507>.
- Reigado, J., Rodrigues, H., 2017. Vocalizations produced in the second year of life in response to speaking and singing. *Psychology of Music* 1–12. <https://doi.org/10.1177/0305735617719335>.
- Reissland, N., Stephenson, T., 1999. Turn-taking in early vocal interaction: A comparison of premature and term infants' vocal interaction with their mothers. *Child: Care, Health and Development* 26 (6), 447–456.
- Rosenthal, M., 1982. Vocal dialogues in the neonatal period. *Developmental Psychology* 18 (1), 17–21.
- Salerni, N., Suttora, C., D'Odorico, L., 2007. A comparison of characteristics of early communication exchanges in mother-preterm and mother-full-term infant dyads. *First Language* 27 (4), 329–346. <https://doi.org/10.1177/0142723707081654>.
- Stern, D.N., 1985. *The interpersonal world of the infant: A view from psychoanalysis and developmental psychology*. Basic Books, New York.
- Stern D.N., 1999. Vitality contours: The temporal contour of feelings as a basic unit for constructing the infant's social experience. In: Rochat, P. (Ed.), *Early social cognition: Understanding others in the first months of life*. Erlbaum, Mahwah, NJ, pp. 67–90.
- Stern D.N., Hofer, L., Haft, W., Dore, J., 1985. Affect attunement: The sharing of feeling states between mother and infant by means of inter-modal fluency. In: Field, T.M., Fox, N.A. (Eds.), *Social perception in infants*. Ablex, Norwood, N.J., pp. 249–268.
- Striano, T., Henning, A., Stahl, D., 2005. Sensitivity to social contingencies between 1 and 3 months of age. *Developmental Science* 8, 509–518. <https://doi.org/10.1111/j.1467-7687.2005.00442.x>.
- Trainor, L.J., Adams, B., 2000. Infants' and adults' use of duration and intensity cues in the segmentation of tone patterns. *Perception & Psychophysics* 62 (2), 333–340.
- Trevarthen, C., 1993. The self-born in intersubjectivity: The psychology of an infant communicating. In: Neisser, U. (Ed.), *The perceived self: Ecological and interpersonal sources of self-knowledge*. Cambridge University Press, New York, NY, pp. 121–173. <https://doi.org/10.1017/CBO9780511664007.009>.
- Vannasing, P., Florea, O., González-Frankenberger, B., Tremblay, J., Paquette, N., Safi, D., et al., 2016. Distinct hemispheric specializations for native and non-native languages in one-day-old newborns identified by fNIRS. *Neuropsychologia* 84, 63–69. <https://doi.org/10.1016/j.neuropsychologia.2016.01.038>.
- Van Egeren, L.A., Barratt, M.S., Roach, M., 2001. Mother–infant responsiveness: Timing, mutual regulation, and interactional context. *Developmental Psychology* 37, 684–697. <https://doi.org/10.1037/0012-1649.37.5.684>.